Disparity map coding for 3D teleconferencing applications

Article *in* Proceedings of SPIE - The International Society for Optical Engineering · March 1997

DOI: 10.1117/12.263249 · Source: CiteSeer

2 authors:

André Redert
Rodotti

Emile Hendriks
Delft University of Technology

Disparity map coding for 3D teleconferencing applications

André Redert, Emile Hendriks

Information Theory Group, Department of Electrical Engineering
Delft University of Technology,
Mekelweg 4, 2628 CD Delft, The Netherlands
email: andre,emile@it.et.tudelft.nl,
phone: +31 15 278 6269, fax: +31 15 278 1843

ABSTRACT

In this paper we present a combination of three steps to code a disparity map for 3D teleconferencing applications. First we introduce a new disparity map format, the chain map, which has a very low inherent redundancy. Additional advantages of this map are: one single bidirectional map in stead of the usual two unidirectional vector fields, explicit indication of occlusions, no upper or lower bound on disparity values, no disparity offset, easy generation by disparity estimators and easy interpretation by image interpolators. In a second step, we apply data reduction on the chain map. The reduction is a factor two, thereby losing explicit information about the position of occlusion areas. An algorithm for image interpolation in absence of occlusion information is presented. The third step involves entropy coding, both lossless and lossy. A scheme specially suited for the chain map has been developed. Although the codec is based on a simple prediction process without motion compensation, compression ratios of 20 (lossless) to 80 (lossy) can be achieved with typical teleconferencing images. These results are comparable to those obtained by complex schemes based on 2D/3D motion compensation using disparity vector fields.

Keywords: disparity map coding, disparity map format, 3D teleconferencing, image interpolation

1. INTRODUCTION

In teleconferencing, the telepresence feeling is enhanced substantially by the introduction of 3D images. In addition to stereo teleconferencing, this implies the need for disparity estimation[2-6] and image interpolation[1,3,5,7,9]. We assume that disparity estimation is performed at the transmitter and image interpolation at the receiver. This allows for a low complexity receiver in multiperson communications, but requires the coding and transmission of the disparity map. Disparity map coding is still a relatively new area. Tzovaras et al.[10] provide an overview of different coding methods for disparity vector fields.

In this paper we will examine a combination of three steps to code a disparity map. First we present a new disparity map format, the chain map, and examine its properties and redundancy. Secondly, we will apply data reduction on the chain map, loosing explicit occlusion information. A new image interpolation algorithm is presented that does not need this information. The reduction also lowers the complexity of the disparity estimator at the transmitter. Finally we investigate both lossless and lossy entropy coding using schemes especially suited for the chain map.

2. DISPARITY MAP CONSTRAINTS

A disparity map describes which pixels in stereo images form pairs describing the same 3D scene point. Our constraints on the disparity map are:

- **Disparity vectors have a horizontal component only.** All pixels in one scanline of the left image correspond only to right image pixels on the same scanline. To achieve this, a stereo camera setup with parallel optical axes is sufficient. In the case of a converging camera set-up, a rectification process can be applied[8].
- **The ordering constraint is obeyed.** This means that the from-left-to-right-order of objects is the same in the left and right images. The ordering constraint may be violated when scenes with small objects at different depths are recorded by two camera's with large baseline. This might occur in teleconferencing scenes by exception (a person reaches out his or her hand in front of the camera's). Many disparity estimation algorithms[2,6,7] need the ordering constraint, in which case our constraint is not additional.

- **In occluded areas, the disparity map carries no pseudo disparity information.** When an object or part of it is visible in only one of the images, accompanying disparity vectors do not exist. Still one can generate pseudo disparity vectors, pointing from the visible part in one image to the imaginary counterpart in the other image. For this one needs sophisticated algorithms that exploit monocular depth cues, or more camera's that provide additional depth information. In many cases, however, pseudo disparity vectors are generated in a postprocessing step using a linear or constant fill operation on the calculated real disparity vectors[1,3,4,5]. Since the fill operation provides no additional information, we generate the pseudo vectors during image interpolation and exclude their presence in the disparity map.

3. A NEW DISPARITY MAP FORMAT: THE CHAIN MAP

A well known format for disparity maps is the disparity vector field[1-10], sometimes explicitly accompanied by occlusion labels[1]. Although used very commonly, the vector field format has the following disadvantage.

Scene object surfaces that are normal to the camera axes result in projections of the same size in left and right image. We call this normal objects. When an object surface is not normal to the camera axes, the projections in left and right images have different size. We call this contractions. We define a left contraction as an object that is larger in the left image than in the right image. An object that is visible in the left image, but not in the right image, is defined as a left occlusion.

A single right-to-left vector field is unable to indicate the difference between left occlusion and a strong left contraction. In both cases the disparity vectors are diverging, leaving pixels in the left image unpaired. This problem can be solved by using two fields[1,7,9], but that introduces much redundancy.

In stead of vector fields we propose to use a new format, the chain map. Figure 1 shows the chain map format of one scanline of a disparity map. For clearness, the left figure shows the disparity vector representation[1,3,9] and the right figure shows the disparity path representation[2,3,6,9].

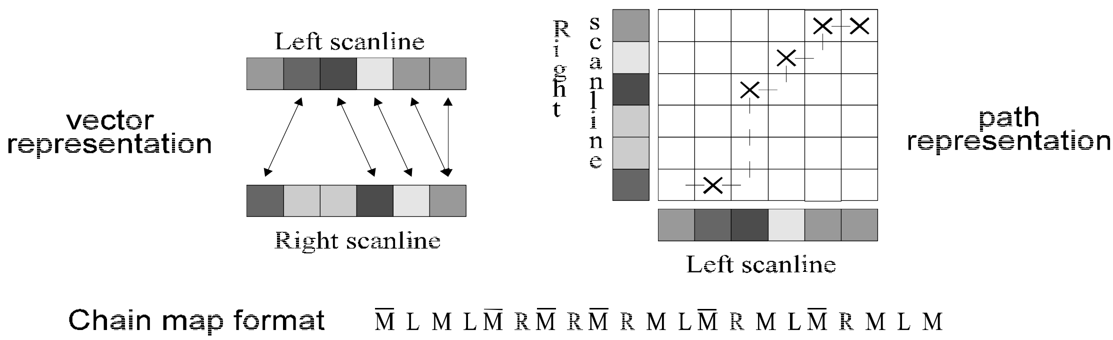

Figure 1: The chain map format of a disparity map

Starting at the first pixel on corresponding left and right image scanlines, we indicate whether a disparity vector is present between the two pixels (M, a matching pair) or not ($\overline{M}$). Then we take a 1-pixel horizontal step in one of the two scanlines (L or R) and again indicate whether a vector is present, etc. This continues until the edge in both scanlines is reached.

All possible disparity maps under our constraints fit in the chain map. Occlusions and objects are indicated explicitly by the chain map. The match bits M indicate normal or contracted objects. In the case of contracted objects it is possible that single pixels are part of more than one pair. The match bits $\overline{M}$ indicate occlusions.

The step bits L/R have the following meaning. If left image pixel xl,y and right image pixel xr,y form a disparity pair, we define disparity as $d(x) = xl - xr$, with $x = xl + xr$. The pixel co-ordinates x,y correspond to the image of an imaginary centre camera, that has double horizontal size. Now we define S as the continuous derivative of d with respect to x, and $S^{\#}$ as the discrete derivative, valued either -1 or +1:

$$S(x) = \frac{\partial d}{\partial x} \quad , \quad S^{\#}(x) = \frac{\partial^{\#}}{\partial x} d \qquad (1)$$

The ordering constraint, described in section 2, together with the definition of d and x, ensures that $|S| \leq 1$. Therefore the discrete derivative operator does not become 'saturated' (happens when $|S| \geq 1$ while $|S^{\#}| = 1$) and the integral of $S^{\#}$ resembles d closely. On the other hand, S and $S^{\#}$ do not resemble each other on local scale, since the quantisation noise in $S^{\#}$ has a magnitude of the same order as the signal S that is present in $S^{\#}$.

Each L step bit increments d by one, and each R step decrements d by one. Both steps increment x by one. The step bits indexed by x are equal to $S^{\#}(x)$.

A constant disparity results in an alternating series of +1 (L) and -1 (R) step bits. At pixel scale, the discrete derivative is very non linear. On a slightly larger scale, the operator is linear in the sense that the integral of $S^{\#}$ resembles d closely. Therefore linear filtering like smoothing and interpolation can be performed easily on chain maps. Filtering in this domain is equal to permutation of the step bits.

Advantages of the chain map are

- One bidirectional disparity map
- Explicit indication of occluded, contracted and normal objects
- No upper/lower bounds or offsets for disparity values
- Easy generation by estimators that use dynamic programming, close connection with disparity path representation
- Easy interpretation by interpolators

In the next sections we will analyse the redundancy of the step bits of the chain map and provide means to discard the match bits.

4. REDUNDANCY OF THE CHAIN MAP

In this section we will examine the redundancy of the chain map step bits. The results will be used in the design of the entropy codec in section 6.

We divide the analysis in three parts. First we examine the redundancy inherent to the chain map format, without using a priori knowledge about the scene disparity. Then we analyse the case of bounded disparity range, caused by a particular camera setup and/or global a priori knowledge about the scene. Finally we assume that disparity is locally smooth within objects.

4.1 Redundancy inherent to format

If N is the number of pixels on one scanline, then every scanline of the chain map consists of $2N$-1 match bits and $2N$-2 step bits. The total number of bits required for one scanline is thus $4N$-3, which is about 4 bits per pixel measured in original left or right image size. Because the number of L's and R's are both N-1, the total number of possible chain maps is:

$$\frac{(2N-2)!}{(N-1)!^2}\cdot 2^{2N-1}$$

Taking the ^{2}log gives the minimum number B of bits required to code a chain map. Table 1 gives the redundancy 1-B/(4N-3) for different N, obtained using Stirling's approximation formula:

$$\ln x! \approx (x+\frac{1}{2})\ln x - x + \frac{1}{2}\ln 2\pi + \frac{1}{12x} + O(\frac{1}{x^2}) \qquad (2)$$

N	**4*N*-3**	***B***	**Redundancy**
10	37	35	0.054
100	397	393	0.010
256	1023	1017	0.0039
720	2877	2872	0.0017

Table 1: The inherent redundancy of the chain map

For the international standard image format CCIR601/656 N equals 720 and the inherent redundancy of the chain map is less than 0.2 %.

4.2 Bounded disparity range

When we have a priori knowledge about the camera setup and/or scene disparity, we can obtain upper and lower bounds for the disparity values. In the case of a parallel camera setup without shifted lenses, disparity will be always positive. The chain map has no bounds by nature, which increases the relative redundancy. Table 2 gives the results for common values of *N* and upper/lower bounds. The results are obtained by an exact calculation of the number of possible chain maps, using 1500 bits precision integers.

N	**Range**	**4*N*-3**	***B***	**Redundancy**
256	0 ... 256	1023	1009	0.014
256	100 ... 200	1023	810	0.208
256	10 ... 20	1023	972	0.050
256	-5 ... +5	1023	994	0.028
720	0 ... 720	2877	2862	0.005
720	100 ... 300	2877	2663	0.074
720	10 ... 30	2877	2828	0.017
720	-10 ... +10	2877	2854	0.008

Table 2: The redundancy of the chain map, bounded disparity range

Clearly, the redundancy is very low as long as zero disparity is within the allowed range, even for very small disparity ranges. When zero is not included in the range, the redundancy may become quit large. With $Q = |d|_{\min}$, the chain map will start and end with a series of *Q* equal step bits and *Q* match bits $\overline{\mathrm{M}}$. This is easily removed by an external codec.

4.3 Object surface smoothness

Within object surfaces, we assume disparity to be smooth in both the spatial and temporal domain. For every point in the surface we can linearise the disparity in a small neighbourhood. Taking the origin as an example, we define:

$$d(x,y,t) = o + px + qy + rt + f(x,y,t) \tag{3}$$

Here *d* is the disparity of pixel *x*, *y* at time *t* in the centre image. The variables *o*, *p* and *q* correspond to the depth and orientation of the object surface at *t*=0. The variable *r* corresponds to object motion orthogonal to the line between object and optical centre of an imaginary centre camera (which is normal motion in centre image). Just as in optical flow calculations, *r* is a function of *p*, *q* and centre image motion vectors. The time function *f* represents object motion towards or from the centre camera optical centre's. Combining (1) and (3) gives:

$$S(x,y,t) = p + \frac{\partial f}{\partial x} \tag{4}$$

So within object surfaces, *S* provides information about horizontal orientation of that surface and horizontal differences in motion towards or from the camera optical centre. The latter corresponds to object rotation in a scene plane that is projected onto one scanline in the left, right and centre images.

Effectively, due to the missing *r*, taking the continuous derivative *S* of the disparity field *d* incorporates some form of first order motion compensated coding. This argument holds also for the chain map step bits $S^{\#}$. According to (4) we can predict *S* very easily from *S* values in a local spatial or temporal neighbourhood, due to the constant *p*. This is not the case for $S^{\#}$, since the signal to noise ratio is about one, see section 3. Although the local error in predicting $S^{\#}$ step bits is large, the error measured by the integral of $S^{\#}$ and *d* can be very small. So for lossy coding a predictor for single step bits might prove useful.

A single step bit predictor can also be useful for lossless coding. This is the case when the estimator returns a single disparity vector field, which is transformed into a chain map by postprocessing. Between two equal consecutive disparity vectors both LR and RL step bits can be present in the chain map, but the postprocessor might use only one of the combinations.

At object edges the linearisation (3) is not valid. We expect occlusion areas there, corresponding to $|S| = |S^{\#}| = 1$. This results in relatively long runs of equal step bits in $S^{\#}$. These are very easily predictable using a spatial predictor. The motion compensation argument does not hold when (3) is not valid, so a temporal predictor performs better in combination with motion compensation.

Concluding, the chain map exhibits some form of first order motion compensated coding within object surfaces. For those surfaces, exact prediction of a single step bit of $S^{\#}$ is difficult due to the low signal to quantisation noise ratio of $S^{\#}$. However, spatial prediction of step bits is possible for lossy coding and in some cases, dependent on the disparity estimator, for lossless coding. At object edges, also a (motion compensated) temporal predictor can be applied.

5. IMAGE INTERPOLATION USING A REDUCED CHAIN MAP

In this section we will present an image interpolation scheme that does not need the match bits of the chain map. Therefore, a data reduction of a factor two is obtained. The remaining number of bits in the reduced chain map is $2N$-2, slightly below 2 bits per pixel.

Using the reduced chain map it is not possible to distinguish occlusions from contractions explicitly. However, image interpolation algorithms normally need this information[1,3,9,10]. For the luminance I_M of a pixel in the interpolated image, most interpolation algorithms use image data from either left or right image in occlusion areas and a weighted average of both images in object areas:

$$I_M = W_L \cdot I_L + W_R \cdot I_R \qquad \text{with } W_L + W_R = 1 \tag{5}$$

In left occlusions $W_R = 0$, in right occlusions $W_R = 1$ and in object areas $W_R = 1/2 + P$, with P equal to the position of the virtual intermediate camera:

$$\text{(LEFT)} \quad -\frac{1}{2} \le P \le +\frac{1}{2} \quad \text{(RIGHT)} \tag{6}$$

An extension to this is to allow for a local gradual change in W_R at transitions between object and occlusions to avoid luminance discontinuities[7].

We propose a new algorithm that generates similar weights based on the reduced chain map, without using explicitly object-occlusion transition information. Since the chain map does not provide pseudo disparity vectors, the image interpolation algorithm has to calculate them. We use the linear fill method, because in this case occlusion areas are handled as contractions.

For the calculation of the luminance weights W_R and W_L we proceed as follows. First we determine the number N_R of R steps in a symmetric window of length L_{window} around each match bit position. We define

$$State = 2\frac{N_R}{L_{window}} - 1 \tag{7}$$

In a left occlusion or very strong left contraction, $State$ = -1, and in a right occlusion or very strong right contraction, $State$ = + 1. We calculate the weight W_R according to:

$$W_R = W + \Delta W \cdot State \tag{8}$$

and

$$W = \frac{1}{2} + P \quad , \quad \Delta W = \frac{1}{2} - |P| \tag{9}$$

The ΔW part accounts for the automatic weight adaptation according to contraction and occlusion areas. Due to the non zero window length L_{window}, , the small scale quantisation noise of the step bits $S^{\#}$ is removed and gradual weight transitions are provided avoiding luminance discontinuities[7].

When the reduction of the chain map is performed within the disparity estimation process, the complexity of the estimator decreases substantially. For illustration, figure 1 shows the disparity path obtained by a typical estimator using dynamic programming[2,3]. In chain map terms, the middle three matches of the disparity path enclose two occlusions (indicated by two $\overline{\mathrm{M}}$ bits at the end of the chain map). A more plausible interpretation would be that the matches represent a continuous path. To distinguish between these small holes and real occlusions, some threshold on the size of the hole is necessary[3]. With the presented interpolation algorithm and reduced chain map this decision step is not needed.

6. ENTROPY CODING

The disparity maps of natural scenes are a small subset of all possible maps, so entropy coding can be used for compression. Disparity map coding is still a relatively new area. Tzovaras et al.[10] provide an overview of different coding methods for disparity vector fields. We will examine both lossless and lossy compression schemes, specially suited for the reduced chain map.

Figure 2 shows a standard coding system in which we replaced the quantiser by a permutator, explained in section 6.2. Our approach is based on spatial prediction of step bits, thus without using motion compensation.

The coder operates in the following way. Based on all previous reconstructed step bits S^R , a prediction S^P is obtained for the current source step bit S. The prediction error Δ^P can be either -2, 0 or +2. The real prediction errors Δ^P are transformed by the permutator to lossy prediction errors Δ^T, to be transmitted. Based on S^P and Δ^T , the reconstructed step bits S^R are obtained. In the lossless case $\Delta^T = \Delta^P$ and $S^R = S$.

For transmission of non zero Δ^T, the sign of Δ^T can be recovered from S^P, available at the decoder. So the transmitted error signal has a binary form, hopefully including many zero's. A lossless adaptive Huffman codec is used, with input symbols equal to the lengths of zero-runs.

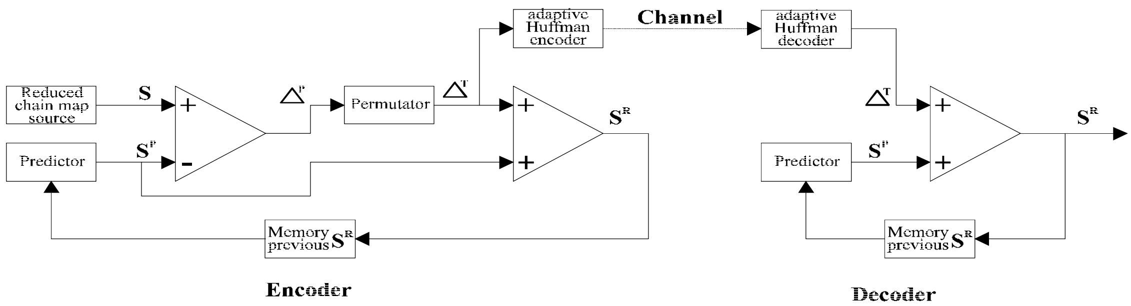

Figure 2: The coding system

6.1 Lossless coding

As predictor we use a simple spatial predictor that takes the step bit from the previous scanline:

$$S^P(x, y, t) = S^R(x, y-1, t) \tag{10}$$

For the top scanline, we take the top scanline of the previous frame. For the top scanline of the first frame we take the alternating step bit series L R L R etc.

This predictor handles the redundancy due to a bounded disparity range, discussed in section 4.2, and the occlusion area and estimator redundancy, discussed in section 4.3.

In this case the permutator is not present, so $\Delta^T = \Delta^P$ and $S^R = S$.

6.2 Lossy coding

For lossy coding in chain maps, we use a combination of two approaches. The first is a K-fold vertical subsampling of the reduced chain map as a preprocessing step. Since many block based algorithms[3,4] create a full size disparity map by interpolating a sparse map, the degradation due to the codec subsampling/interpolation processes may be very small.

Secondly, we add the permutator to the lossless system. The effect of the permutator is that one scanline of S^R is globally the same as S, but with groups of step bits permuted locally. The loss L due to the permutator is defined as:

$$L = \Delta^T - \Delta^P \tag{11}$$

For the reconstruction error Δ^R it follows from figure 2 that:

$$\Delta^R = S^R - S = L \tag{12}$$

Our permutator algorithm works as follows. First we introduce a runlength string notation for one scanline of the Δ^P and Δ^T errors:

$$\{\ 0,0,0,-2,0,+2,0,0,0,+2,+2,0,0\ \} \quad \Leftrightarrow \quad \{\ 3\downarrow 1\uparrow 3\uparrow 0\uparrow 2\ \} \tag{13}$$

The arrows between the runlengths indicates the sign of the nonzero error. Our algorithm is based on the following two lossy transitions:

$$\left.\begin{matrix}(a\downarrow n\uparrow b)\\(a\uparrow n\downarrow b)\end{matrix}\right\}\Rightarrow(a+n+b+2) \tag{14}$$

In the runlength string we search for groups of three runlengths and two opposite arrows, using a specific n. Then we replace the group by one number, thereby replacing three runs by one. Table 3 shows an example with n = 2. Note that the prediction errors in step bits 2 and 4 are not transmitted and thus not corrected, while the prediction errors in step bits 6 and 7 are corrected. In the end, comparing S and S^R , we see that step bits 2 and 4 have been permuted. This has only local influence on the disparity values.

Step bit	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
$\boldsymbol{S}$	-1	+1	-1	+1	+1	+1	-1	+1
$\boldsymbol{S^P}$	-1	+1	+1	+1	-1	+1	+1	-1
$\boldsymbol{\Delta^P}$	0	0	-2	0	+2	0	-2	+2
$\boldsymbol{\Delta^T}$	0	0	0	0	0	0	-2	+2
$\boldsymbol{S^R}$	-1	+1	+1	+1	-1	+1	-1	+1

Table 3: An example of the lossy coding algorithm with *n*=2

From (1) it follows that the error in the reconstructed disparity d is equal to the integral along x of the error Δ^R in the chain map. A single application of (13) results in a uniform disparity error of -2 or +2 in an interval of length n/2 pixels (there are 2N step bits and N pixels on one scanline).

Our algorithm takes Δ^P as initial runlength string. Then we apply the transition rules, searching for the appropriate groups from left to right. This is done for n between 1 and T_{max} . We start by 1 to ensure that the smallest run transitions, causing the least damage, are performed first. The remaining runlengths string is Δ^T.

7. EXPERIMENTS

In this section we will first discuss the used test sequences and evaluation procedures, followed by our results.

7.1 Test sequences

Figure 3 shows the first frame of the used image sequences MAN, HEAD and AQUA, left position. Table 4 provides additional information.

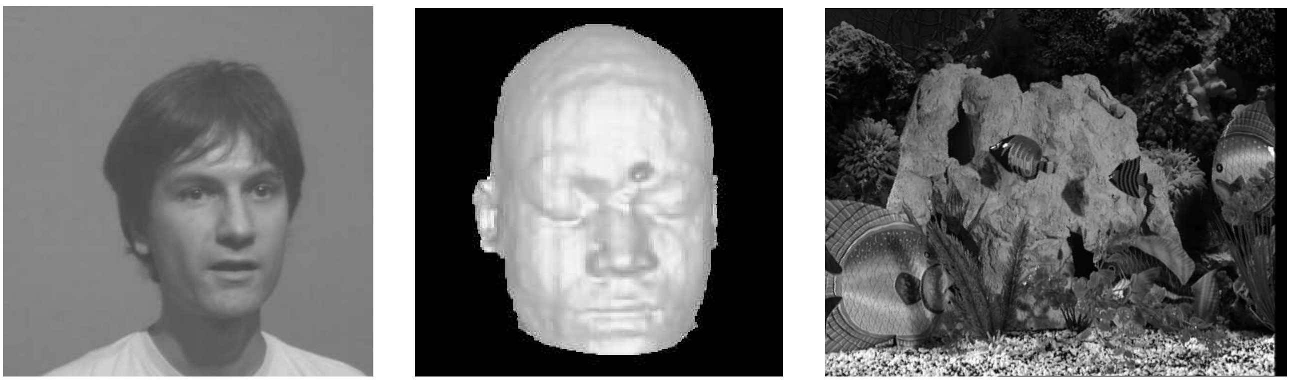

Figure 3: First left frame of MAN, HEAD and AQUA sequence

Sequence name	X*Y size frame	frames/second	Disparity range	Remarks
MAN	384*384	25	30 ... 110	Teleconferencing situation, large camera baseline
HEAD	256*256	25	-5 ... +15	Synthetic medical image, with reference mid position sequence
AQUA	720*576	25	-40 ... +16	RACE DISTIMA test sequence, provided by CCETT France

Table 4: The used test sequences

7.2 Evaluation scheme

To evaluate the original and reconstructed disparity field, we use the following procedure, see figure 4. After estimation of the disparity field D, we generate an intermediate view M' at centre position, based on the original left L and right R image sequence. For the HEAD sequence, M' can be compared to the available original mid position sequence M. For all sequences, we reconstruct left L' and right R' images sequences, performing extrapolation on M' using D.

After that, we code D and after decoding we obtain the reconstructed D*. With D* we repeat the procedure stated above and obtain M*, L* and R*. These can be compared to L, R and M (HEAD sequence) or M' (MAN and AQUA).

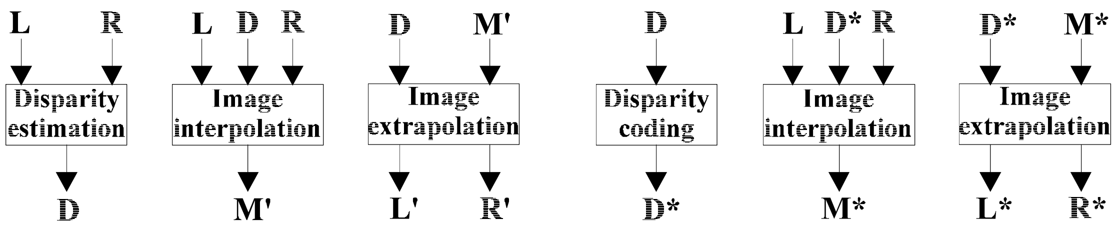

Figure 4: Dependencies between original, reconstructed and coded sequences.

Our disparity estimator is based on the algorithm of Cox et al.[2], p 547, working on blocks of 4*4 pixels rather than single pixels. The algorithm outputs a disparity vector field, with valid vectors at each fourth scanline. Then the vector field is transformed to a reduced chain map and interpolated to a dense map. Figure 5 shows the reduced chain maps, with each luminance value determined by two step bits: RR = black, LR = gray, LL = white. The RL combination is not used, see section 4.3. The disparity range was always set including zero.

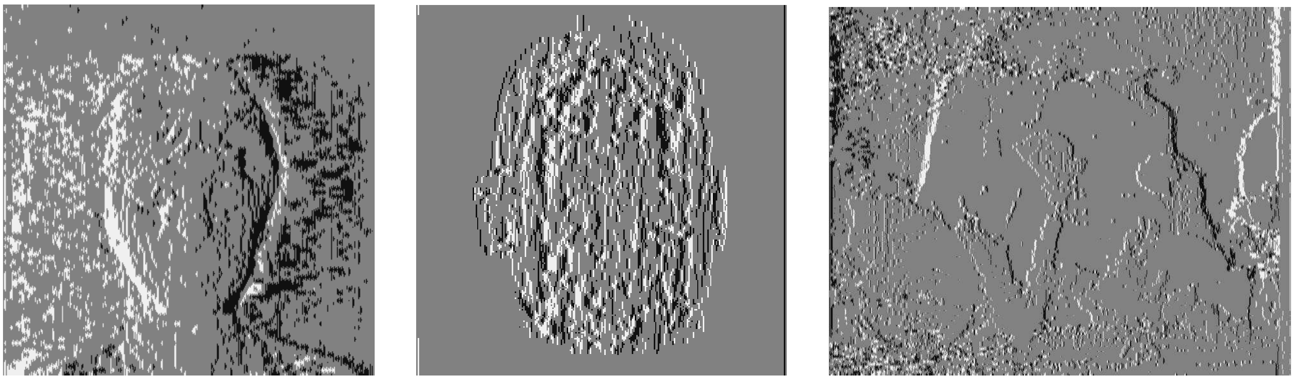

Figure 5: First frame reduced chain maps of MAN, HEAD and AQUA sequence

Our interpolation algorithm is described in section 5. We use L_{window} = 16. Image extrapolation is performed in the same way as interpolation, but without the need to determine interpolation weights (only one source image).

Objective comparison of reconstructed image A and original image B is done using the PSNR with normalisation to maximum luminance I_{MAX} = 255:

$$PSNR = 10 \cdot {}^{10}\log \frac{I_{MAX}^2}{\frac{1}{N_{Pixels}} \sum_{all\ pixels} (I_A - I_B)^2} \qquad (15)$$

7.3 Coding results

All results are based on the first five frames of each test sequence. Table 5 gives the results for lossless coding, with reconstruction PSNR's and bitrates. The bitrates are in bits per pixel and kbits per second. The compression ratio is about 5. The compression is possible due to the spatial redundancy at the occlusion areas (black and white in figure 5) and the redundancy of our estimator, see sections 4.3 and 7.2.

For the lossless case, all degradation of the reconstructed images is due to the disparity estimation and image interpolation processes. Subjectively, all reconstructed images differ only very slightly compared to the original images. The reconstructed MAN sequences look very good, with very small artefacts present only at the ears (occlusion areas). The reconstructed HEAD sequences differ almost unnoticeable from the original sequences. The reconstructed AQUA sequences look good, but slightly smoothed.

Sequence	**PNSR L' / L**	**PSNR M' / M**	**PSNR R' / R**	**bits per pixel**	**kbits per second**
MAN	39.77	-	39.39	0.366	1349
HEAD	32.07	30.19	32.34	0.385	631
AQUA	27.94	-	27.90	0.310	3214

Table 5: Lossless coding results

Table 6 gives the results for lossy coding. Subjectively, we found the quality if the MAN M* image at 0.180 bpp unacceptable, while the other two results look slightly better than M'. This can be explained by the choice of *K*, that matches to the fourfold subsampling of our disparity estimator in the second and third results. On the other hand, all M* result for the HEAD sequence differ only just noticeable from M. The third HEAD results has a compression factor of 300 with respect to an uncoded reduced chain map, and 1200 with respect to an 8 bit disparity vector field. All AQUA results look good, although slightly smoothed as in the lossless case.

Sequence	**T_{MAX} transitions**	***K* vertical subsampling**	**PNSR L* / L**	**PSNR M* / M'**	**PSNR R* / R**	**bits per pixel**	**kbits per second**
MAN	50	1	37.59	42.68	37.65	0.180	663
MAN	10	4	39.20	46.38	38.78	0.128	472
MAN	20	4	38.80	44.89	38.48	0.101	372
HEAD	10	4	30.65	29.88 (to M)	30.59	0.083	136
HEAD	40	8	29.94	29.57 (to M)	29.89	0.018	29
HEAD	80	16	29.28	29.06 (to M)	29.28	0.0067	11
AQUA	10	4	27.47	31.64	27.34	0.087	902
AQUA	20	4	27.31	30.62	27.18	0.060	622
AQUA	40	4	27.02	29.34	26.87	0.037	384

Table 6: Lossy coding results

Figure 6: First frame of MAN sequence, D* (0.101 bpp), M' and M* image

8. CONCLUSIONS

We have presented a combination of three steps to code a disparity map for 3D teleconferencing applications. The first step is the introduction of a new disparity map format, the chain map. We showed that the chain map has very low inherent redundancy. Additional advantages are: one single bidirectional map in stead of the usual two unidirectional vector fields, explicit indication of occlusions, no upper or lower bound on disparity values, no disparity offset, easy generation by disparity estimators and easy interpretation by image interpolators.

In a second step, we applied data reduction on the chain map. The reduction is a factor two, thereby losing explicit information about the position of occlusion areas. We presented an algorithm for automatic occlusion detection, using the reduced chain map.

The third step involves entropy coding, both lossless and lossy. A scheme specially suited for the chain map has been developed. Although the codec is based on a simple spatial prediction process without motion compensation, compression ratios of 20 (lossless) to 80 (lossy) can be achieved for typical teleconferencing images (with reference to an 8 bit disparity vector field). These results are comparable to those obtained by complex schemes based on 2D/3D motion compensation using disparity vector fields.

ACKNOWLEDGEMENT

This work was done in the framework of the European ACTS project PANORAMA. One of the major goals of this project is a hardware realisation of a real time multi viewpoint autostereoscopic communication system. In this system the reduced chain map will be used as disparity map format.

REFERENCES

1. B. Chupeau and P. Salmon, "Synthesis of intermediate pictures for autostereoscopic multiview displays", in Proceedings of the Workshop on HDTV '94, Turin, Italy, 1994

2. I.J. Cox, S.L. Hingorani and S.B. Rao, "A maximum likelihood stereo algorithm", in *Computer Vision and Image Understanding*, Vol. 63, No. 3, 1996, pp. 542-567

3. R.E.H. Franich, "Disparity Estimation in Stereoscopic Digital Images", Ph.D. dissertation, Information Theory Group, Department of Electrical Engineering, Delft University of Technology, 1996

4. E.A. Hendriks and Gy. Marosi, "Recursive disparity estimation algorithm for real time stereoscopic video applications", in Proceedings of the International Conference on Image Processing (ICIP) 1996, pp. 887-890

5. E. Izquierdo and M. Ernst, "Motion/Disparity Analysis and Image Synthesis for 3DTV", in Proceedings of the Workshop on HDTV '94, Turin, Italy, 1994

6. J. Liu and R. Skerjanc, "Stereo and motion correspondence in a sequence of stereo images*", Signal Processing: Image Communication* 5, 1993, pp. 305-318

7. J. Liu, I.P. Beldie and M. Wöpking, "A computational approach to establish eye-contact in videocommunication", in Proceedings of the International Workshop on Stereoscopic and Three Dimensional Imaging (IWS3DI), Santorini, Greece, 1995, pp. 229-234

8. D.V. Papadimitriou and T.J. Dennis, "Epipolar line estimation and rectification for stereo image pairs", in Proceedings of the International Workshop on Stereoscopic and Three Dimensional Imaging (IWS3DI), Santorini, Greece, 1995, pp. 128-133

9. B.L. Tseng and D. Anastassiou, "A Theoretical Study On An Accurate Reconstruction Of Multiview Images Based On The Viterbi Algorithm", in Proceedings of the International Conference on Image Processing (ICIP) 1995, pp. 378-381

10. D. Tzovaras, N. Grammalidis and M.G. Strintzis, "Disparity field and depth map coding for multiview image sequence compression", in Proceedings of the International Conference on Image Processing (ICIP) 1996, pp. 887-890